International® Baccalaureate

Extended essay cover

Diploma Programme subject in which this extended essay is registered: WORLD STUDIES: SciTech SoS and journal

(For an extended essay in the area of languages, state the language and whether it is group 1 or group 2.)

Title of the extended essay: Modelling the Future of Biochemistry Research

Candidate's declaration

This declaration must be signed by the candidate; otherwise a grade may not be issued.

The extended essay I am submitting is my own work (apart from guidance allowed by the International Baccalaureate).

I have acknowledged each use of the words, graphics or ideas of another person, whether written, oral or visual.

I am aware that the word limit for all extended essays is 4000 words and that examiners are not required to read beyond this limit.

This is the final version of my extended essay.

Supervisor's report and declaration

The supervisor must complete this report, sign the declaration and then give the final version of the extended essay, with this cover attached, to the Diploma Programme coordinator.

Name of supervisor (CAPITAL letters)

Please comment, as appropriate, on the candidate's performance, the context in which the candidate undertook the research for the extended essay, any difficulties encountered and how these were overcome (see page 13 of the extended essay guide). The concluding interview (viva voce) may provide useful information. These comments can help the examiner award a level for criterion K (holistic judgment). Do not comment on any adverse personal circumstances that may have affected the candidate. If the amount of time spent with the candidate was zero, you must explain this, in particular how it was then possible to authenticate the essay as the candidate's own work. You may attach an additional sheet if there is insufficient space here.

This declaration must be signed by the supervisor; otherwise a grade may not be issued.

I have read the final version of the extended essay that will be submitted to the examiner.

To the best of my knowledge, the extended essay is the authentic work of the candidate.

I spent [4] hours with the candidate discussing the progress of the extended essay.

Assessment form (for examiner use only)

Criteria	Achievement level				
	Examiner 1	maximum	Examiner 2	maximum	Examiner 3
A research question	1	2		2	
B introduction	1	2		2	
C investigation	2	4		4	
D knowledge and understanding	3	4		4	
E reasoned argument	2	4		4	
F analysis and evaluation	3	4		4	
G use of subject language	3	4		4	
H conclusion	1	2		2	
I formal presentation	2	4		4	
J abstract	1	2		2	
K holistic judgment	2	4		4	
Total out of 36	21				

Modelling the Future of Biochemistry Research

A World Studies Extended Essay studying the evolution of Molecular Modelling techniques

By:

Abstract:

The purpose of this essay is to study how molecular models have evolved to better serve medical research and how they are giving scientists a better picture of the structure and function of the molecules they analyze. This topic is globally significant since it relates to the issue of medical research and for that same reason is locally significant, since every community in the world can benefit from improved medical research and treatments. This essay will focus on previous molecular modelling techniques, such as chemical drawings, sequence diagrams and three-dimensional (3D) plastic models. It will then discuss current techniques that focus on computerized 3D modelling, as well as 3D printouts. Finally, it will suggest future improvements that could be made to molecular modelling techniques in order to further assist biochemical and medical research.

My interest in this topic began when I joined my high school's SMART (Student Modelling A Research Topic) Team. Throughout my years in the program, we have studied various molecular models and examples from an enzyme that is partly responsible for allergic reactions to bee stings, the activator for the bubonic plague, an enzyme that can lead to stomach ulcers, a protein that is involved with SARS, and an enzyme that when expressed, can lead to cancer. I have had the opportunity to work on a variety of molecules, and it has opened my eyes to the immense amount of uses and advantages of new molecular modelling techniques.

Word Count: 3864

Acknowledgements:

I would like to acknowledge the following sources that I consulted in the initial stages of my research for this essay.

Fukuda, M., Rutishauser, U., & Schnaar, R. L. (Eds.). (2005). *Neuroglycobiology.* Oxford: Oxford University Press.

Goodsell, D. S. (2009). *The machinery of life* (2nd ed.). New York, N.Y.: Copernicus Books.

Petsko, G. A., & Ringe, D. (2004). *Protein structure and function.* London: New Science Press.

Rao, F. V., Rich, J. R., Rakiu0107, B., Buddai, S., Schwartz, M. F., Johnson, K., . . . Strynadka, N. C. J. (2009). Structural insight into mammalian sialyltransferases [PDF]. *Nature Structural and Molecular Biology*, 1-3.

Rasheed, M., Elmore, S. C., & Wirth, T. (n.d.). 8.2.7 Selective Glycosylation. In M. Gruttadauria & F. Giacalone (Eds.), *Catalytic Methods in Asymmetric Synthesis* (pp. 352-353).

Table of Contents:

The fields of biology and chemistry have been related to each other since the beginning of their study, with biochemistry being the result. In recent years, all of the sciences, specifically biochemistry and genomic sciences, are beginning to become more intertwined with the field of computer science (Field, 2007; Carbajo & Tramontano, 2012). In today's society, as technology develops and improves, so does science, as our understanding of nature and specifically of molecules broadens. This essay is an interdisciplinary approach because in reality any study in the sciences is now an interdisciplinary study, since one cannot talk about science without mentioning the technology involved. One of the fastest growing applications of both biochemistry and computer science is that of molecular modelling. Molecular modelling techniques are constantly evolving, as are computers.

The evolution of molecular models is an issue of contemporary global significance. Often, science is dealt the task of solving a problem in the world, whether that be climate change or finding a cure to cancer. The purpose of this essay is to study how molecular models have evolved to better serve medical research and how they are giving scientists a better picture of the structure and function of the molecules that they analyze. This topic is globally significant since it relates to the issue of medical research and for that same reason is locally significant, since every community in the world can benefit from improved medical research and treatments.

This topic has also gained importance because of the fact that drug companies are focusing much more of their efforts on so-called lifestyle drugs rather then on drugs to treat heart disease or cancer, since there is more money to be made in lifestyle drugs (Laurance, 2012; Thompson, 2011). Molecular models can be used in the field of

medicine to help scientists find novel approaches to diseases and viruses, and ideally help them create new drugs to treat illnesses (Burwash, Zhang, Eid-Ricci & Snoddy, 2013). If it becomes easier, and therefore more cost effective, for companies to research traditional illnesses, then they might be more motivated to focus on traditional drugs. This is especially true when it comes to the reactions catalyzed by enzymes in the body that can lead to illnesses. For example, scientists, through the use of various models to see which part of the enzyme is the active site, can find a way to stop that portion of the enzyme (Burwash et al., 2013). Using newer three-dimensional (3D) modelling techniques, scientists and researchers can study the exact shape of the active site. With this knowledge, it is possible to figure out methods of both competitive and non-competitive inhibition that could be used to stop that enzyme from performing the reaction that causes the illness. Essentially, the better the quality, detail, and accuracy of the model, the higher the likelihood that scientists will be able to stop many of the diseases that exist.

Before discussing molecular models and the ways in which they have evolved, it is important to first understand the word model as defined in science. Hinchliffe uses Chambers Dictionary's definition of a model, which states that a model in regards to science is a "plan, design; a preliminary solid representation generally small, or in plastic material, to be followed in construction; something to be copied: a pattern: an imitation of something on a smaller scale: a person or thing closely resembling another" (Hinchliffe, 2003, p.5). This definition appears vague due to the large number of variants that it includes, but this is because of the nature of models in science. Almost anything can be used as a useful model, and there are an enormous amount of forms that these models could take.

It is also important to note that although models are used widely throughout science, they are often 'inaccurate', each type of model has pros and cons. Schlick applies a quotation from Pablo Picasso, stating that "Art is the lie that helps tell the truth" (Schlick, 2002, p.3) and states that the same can be said for models. The idea of a scientific model is that it represents something that exists in the real world, it is not meant to be exact, it is simply supposed to allow the researcher to view the molecule in an alternate mindset. Also, there are often many different models of the same thing, even for something as simple as water. Each of these different models serves a purpose, and will be used by scientists studying different aspects of the item.

As for molecular modelling, it is "the science and art of studying molecular structure and function through model building and computation. The model building can be as simple as plastic templates or metal rods or as sophisticated as interactive animated colour stereo graphics and laser-made wooden sculptures." (Schlick, 2002, p.3). It is stated that molecular models, like all other scientific models, can come in a magnitude of forms ranging from the simple to the extremely complex. The limitations of molecular models currently limit the understanding of molecules.

To successfully study how molecular modelling techniques have evolved, it is vital to look at previous techniques, some of which are still used today alongside newer techniques. The first of these previous techniques is that of text based descriptions. These come from the name of the molecule being studied. They can be useful for basic compounds, as well as organic compounds that follow the nomenclature rules of organic chemistry. This technique is beneficial since the only "technology" that it requires is a basic understanding of organic chemistry. However, this method becomes much less

valuable when the names of the compounds become more complicated, or if they are known by other, common names. It also tends to be less effective when looking at complex proteins, which do not follow the same specific naming rules as those of organic chemistry.

The second category of previous techniques is that of two-dimensional (2D) models. There are many different types of 2D models used by scientists in the field of biochemistry. Chemical drawings (see Appendix I Figure 1) are an example of 2D molecular model. The main advantages of these molecular models are that they can be drawn very quickly and can be created simply and cheaply. An example of this would be the chemical drawing of water, shown earlier in this paper. They are rarely used in higher level research because of their many disadvantages such as giving the impression that all of the molecules that are being studied are linear and 2D. If someone wants to represent a molecule or compound in 3D using chemical drawings, they are forced to make at least six drawings, showing the molecule from multiple viewpoints (Hinchliffe, 2003). This can help those studying the diagram to understand the 3D molecule from 2D drawings, but becomes vary labour intensive for more complicated molecules. There are computer programs that are designed to create 2D chemical drawings. They do save some time because of the fact that they can create the drawings faster than the human hand, but still encounter many of the same difficulties.

Another example of a 2D model, and one that is primarily used for more complex protein molecules is that of a sequence diagram (see Appendix I Figure 2). These are the most complicated 2D model types, using letters to represent the different amino acids that make up the peptide. These diagrams, like all types of models, do have some advantages,

however, there are many more disadvantages. The disadvantage with this type of model is that they only show the primary protein structure of the molecule, and therefore give no sense of the molecule as it exists in three dimensions. They can also be very complicated to interpret without the use of a legend that explains which amino acid is represented by which letter.

The final previous technique for molecular modelling is that of 3D manual models. These are the typical ball and stick plastic models that are used by students to build simple models in chemistry classes. The benefits of these models are that they are relatively cheap and simple to build. They are also beneficial since they are built in 3D and show the relative shape and bond angles of the atoms involved in the molecule/compound. The main problem with these models is that they are extremely time consuming, especially for large molecules such as proteins that can have thousands of atoms. The task for a scientist to manually create this model is immense, and therefore it becomes extremely impractical. The other problem with this type of molecular modelling is that it requires the model maker to know the exact structure of the molecule, whereas newer technologies allow for work to be shared and allow researchers to study part of a molecule without knowing its entire structure (Hinchliffe, 2003). This type of modelling was quite prominent around the 1970s, but nowadays is used more in the classroom setting than in the research labs.

With these previous techniques explained, it is now possible to see how they have evolved into the molecular modelling techniques used today by researchers in the medical fields (see Appendix I Figures 3 and 4). The current method for molecular modelling uses computer technology, and usually involves some sort of command line programming

(See Appendix II). To make these models, the macromolecule is typically crystallized, a process which is neither cheap nor easy. Using the information from the crystallization, a digital file can be created, typically through command line programming but depending on the program, (for example, the Java based Jmol uses user inputted simplified Java commands). At first creation, these types of models could only be made on supercomputers designed specifically for the task of molecular modelling due to the large amount of processing power required. However, nowadays, thanks to the improvements that have been made to computer processors, molecular modelling programs can be run on almost any computer and in some cases mobile devices (1.14 Hardware and Software, 2012). One example of these modern programs is JSmol, which uses JavaScript programming, a lighter version of Java programming designed for mobile devices (Hanson, Prilusky, Renjian, Nakane & Sussman, 2013). The main flaw with the current method of molecular modelling is that the first time that a particular molecule is to be modelled, it must be programmed through command line text. This process is very time consuming, and it is for that reason that there now exist many databases of molecules that have been modelled. These databases include the Protein Data Bank (produced by the Research Collaboratory for Structural Bioinformatics), ModelDB as well as Proteopedia, which allow files of molecules previous modelled to be downloaded by other users (Carbajo & Tramontano, 2012; Hodis et al, 2008). All of these databases used the .pdb file format, which can be read by the majority of molecular modelling programs. The latter of these databases, Proteopedia, has been born out of the evolving uses of the internet. It uses a wiki style webpage that can be edited by registered and approved users.

This is a trend in many research areas, with a shift towards a more collaborative and shared approach.

Once molecular modelling files are downloaded from databases, they can be opened by a variety of molecular modelling programs, many of which are free and open source. These programs allow the molecule to be easily modified to the viewing method chosen, such as wireframe, ball and stick, and spacefill. This allows the scientist to take a standardized file from a database, and then customize it to fit the requirements that they have of the model. These programs also have useful classification systems such as chains, residues and atoms, which are extremely useful when dealing with larger protein molecules (Field, 2007). Due to the growing simplicity and availability of molecular modelling programs, molecular modelling is starting to be used more and more in industrial applications for research. Another advantage to these programs is that they are able to calculate something that would be included on the model, even if it is not present and/or visible. This is an important feature since, if there is a part of the molecule that is in continuous motion, it cannot be crystallized and therefore cannot be added to a molecular model in the normal method. However, using new digital molecular modelling programs and any information that researchers have about the missing portion of the molecule, it is possible to digitally recreate the missing section so that it can be studied on the molecule (Hinchliffe, 2003). This type of reconstruction relies on theories of comparative modelling that state that molecules, specifically proteins, that have evolved the same way will have many of the same structural elements (Carbajo &Tramontano, 2012; Burwash et al., 2013). These models give a sense of motion by allowing the user to manipulate the angle of the image on the screen and zoom in/out as necessary. This is

something that is not offered by any of the early methods; however, they are mostly just 3D snapshots of the molecule. It is important to be able to view the actual motion of a molecule, especially when considering enzymes and their conformational dynamics, as it is important to be able to identify the patterns of motion that are used by the enzyme.

Before discussing how the current digital molecular models could be changed in the future, there is one other new development that is currently being used in the area of molecular modelling (see Appendix I Figure 5). It is the new technology of 3D printing. This process, as it is applied in molecular modelling, uses plastics and binding agents to print models layer by layer. This is an extremely important new development, since detailed models that are created using digital molecular modelling programs can be printed out in a way that encompasses all of their detail, while giving researchers a physical 3D model that they can actually hold in their hands. 3D printing is a technology that is being employed in an increasing number of fields because of its practicality, but also because of the fact that the cost of 3D printing technology is rapidly decreasing. It is now possible to buy a basic 3D printer for only a few thousand dollars; a cost which ends up being relatively small when compared to how much research can be done.

Earlier this year, there was even a 3D Printing Conference in New York City. Mark LePage described it as " you are watching ideas take physical form" (LePage, 2013). Jenifer Howard of MakerBot stated "We've created a machine that allows others to create" (LePage, 2013). Unfortunately, as with many inventions, machines can be used for bad as well as good. Desktop manufacturing, as it is now called, has allowed for the manufacture of guns. This is a controversial result, in fact the US Government is taking steps to cut off public access to these files (LePage, 2013). Discussion surrounding the

use or misuse of 3D printing has brought knowledge of the technique to the general public. Publicity, even bad publicity such as this, may lead to further adoption of 3D printing technology into scientific research.

In the future, there are two main types of improvements that will be made to molecular modelling techniques in order to help them continue to evolve to serve their purposes in medical and biological research. The first of these methods would look at how to make minor improvements to current technologies, while the other would look at creating new technologies.

When it comes to an improved version of current technologies, it would be ideal for scientists to be able to create and view molecular models in a quicker and more efficient manner. The best way to facilitate this would involve the creation and updating of modelling programs and techniques so that they could be used for bigger molecules, which could be made faster, and would be of better quality with more details and accuracy. Many of these improvements will occur naturally, as computers become faster, more efficient and with better internal specifications than at present. The other improvement that could come to existing programs would be to make them much more user friendly. In the case of many of the best molecular modelling programs, they were designed in a way that would allow them to fulfill their purpose, not necessarily be easy to use. One of the biggest improvements would be a user interface that would allow a molecular model to be created (even the first time) without the immense amount of command line programming (Field, 2007). Initially, this would require further development of the programs, but afterward, it would permit researchers to create molecular models without needing the computer science knowledge or background to

perform the command line programming. Other minor improvements for user experience could encompass the compatibility of these programs. This would continue the trend of making these programs available on all major computer platforms (Windows, OS X, etc.) as well as the new trend of making molecular modelling programs available on mobile devices. Cloud computing technology could also be implemented into molecular modelling programs, specifically high-end programs (1.14 Hardware and Software, 2012). This new feature would make the sharing of molecular modelling files easier and would better facilitate multiple users to be working on the same model.

However, the biggest and most important next step in the evolution of molecular modelling techniques is the creation of a new sort of visual video program. This would, in short, give researchers the opportunity to have a movie of their molecule rather than multiple snapshots. This method of modelling would be specifically useful when it comes to the study of enzymes and the reactions that they catalyze. The current theory that is widely used as the definition for enzyme substrate interactions is the Plasmology Theory. This theory states that an enzyme has an ever-changing shape and that it continuously undergoes shape changes to better accommodate its substrate (and ligand if one is involved in the reaction). This theory replaces the previous Induced-Fit model of the enzyme, which worked better with molecular modelling snapshots since it stated that the enzyme would come and wrap around the substrate like a handshake, but would not change its shape. Along with the idea of conformational dynamics, which shows that enzyme's change shape based on the free energy available in the environment, not just the specific substrate. Current modelling techniques simply do not give enough information to be fully useful and correct when looked at from the point of view of

conformational dynamics and the plasmology model. Proteopedia, for example, is starting to incorporate conformational dynamics snapshots into their database, which shows that molecular modelling development is leading in this direction (Hodis et al., 2008).

With current modelling techniques, researchers can only see a snapshot of the enzyme, the substrate and possibly a co-factor (or ligand). It is possible to see them prior to the reaction, and at the end of the reaction, but when it comes to medical research uses, it is important for scientists to be able to study exactly how the substrate enters the enzyme's active site and to see what occurs in the reaction site. The only way for scientists to study how the reaction in the active site occurs is to have a visual video model. Otherwise, they will not be able to study how to inhibit a certain enzymatic reaction, something that is often critical to helping treat many illnesses. If researchers, specifically in the medical fields, want to be able to study why certain reactions occur in the body and thus cause diseases, they have to be able to visualize the way the reaction actually works.

Another interesting addition to future modelling techniques would be the ability to switch between the various domains that make up the molecule in question (1.14 Hardware and Software, 2012). This would, in essence, allow the researcher to view and switch between the extremely detailed atomic and molecular levels but also view a general view of the molecule to show its shape as it exists in the real world.

The term "visual video" is used to describe this new type of molecular model because it could take on multiple different forms. It is unclear exactly how the 3D model would be represented by the video model. Presumably, the molecule would first be

modelled using current 3D modelling techniques, and this step would be repeated at different stages of the reaction process. These multiple snapshots could be then assembled into some sort of film. It is at this point that the idea of how to maintain the 3D aspect has not been thoroughly discussed. It could take the form of a video being played in 3D like many movies, where scientists could view their models on specialized screens. It could also take the form of a 3D projection (almost like a holograph), where the researcher could see a 3D video of their model floating in front of them (Hodis et al., 2008). This version seems quite far off, but could be possible depending on the speed of the evolution of technology in general. Ideally, as video and animation technology become more sophisticated and are used more in the scientific field, then a method could be developed where the computer is able to figure out the form of the reaction without it having to be crystallized at multiple stages of the reaction. This would allow for a full video animation of the molecule, which, due to new development in 3D screens, could be visualized by the researcher in 3D as the reaction occurs. This development would allow researchers to better understand the way that the reactions occur, and therefore be able to treat any diseases, such as cancer, that are caused by these reactions.

Molecular models, in all forms, from something as simple as a chemical drawing, to something as complicated as a future 3D visualization of a reaction, are a crucial part of the modern biological sciences. Models are used constantly today in medical research, and will continue to be used more frequently as their technology improves and becomes more accessible (both from a financial and simplicity to operate point of view). Molecular modelling programs and techniques are starting to become a large industry, and this new industry is continuously growing. These techniques will develop faster as

more companies add molecular modelling tools to their portfolios and the further development of these techniques will hopefully lead to the development of cures to diseases such as cancer. These models have evolved immensely since early versions, and are changing directly with constant improvements in the realm of computer science. For many recent discoveries, biologically and medically, molecular modelling techniques have proven to be both valuable and indispensable. It is amazing to think of how many new discoveries will be made in the future because of the evolution of molecular modelling techniques. This is the era of computers and technology, and with that, it could be said that this will be the generation that watched the evolution of molecular modelling techniques into the mainstream of scientific research.

References:

1.14 Hardware and Software; 1.15 Summary; 1.16 Looking Ahead. (2012). In B. Rai (Ed.), *Molecular modeling for the design of novel performance chemicals and materials* (pp. 20-21). Boca Raton, Fl: CRC Press.

Burwash, A. G., Zhang, B., Eid-Ricci, A., & Snoddy, T. (2013). *Human Sialyltransferase (ST3-Gal1)* [Scientific Poster; Microsoft PowerPoint].

Carbajo, D., & Tramontano, A. (2012). A resource for benchmarking the usefulness of protein structure models. *BMC Bioinformatics*, *13*(188). Retrieved June 18 2013 from www.scopus.com.

Field, M. J. (2007). Chapter 1: Preliminaries and Chapter 2: Chemical models and representations. In *A practical introduction to the simulation of molecular systems* (2nd ed., pp. 1-30). Cambridge: Cambridge University Press.

Hanson, R. M., Prilusky, J., Renjian, Z., Nakane, T., & Sussman, J. L. (2013). JSmol and the next-generation web-based representation of 3D molecular structure as applied to proteopedia. *Israel Journal of Chemistry, 53*(3-4), 207-216. Retrieved June 18 2013 from www.scopus.com

Hinchliffe, A. (2003). Chapter 1: Introduction. In *Molecular modelling for beginners* (pp. 1-11). Chichester, West Sussex, England: Wiley.

Hodis, E., Prilusky, J., Martz, E., Silman, I., Moult, J., & Sussman, J. L. (2008). Proteopedia - A scientific 'wiki' bridging the rift between three-dimensional structure and function of biomacromolecules. *Genome Biology*, *9*(8). Retrieved June 20 2013 from www.scopus.com.

Jmol Quick Reference Sheet [PDF]. (n.d.).

This file is part of the material provided to members of the SMART Team program.

Laurance, J. (2012, August 8). Drugs companies putting profit ahead of medical discoveries, warn scientists. *The Independent*.

Retrieved September 8 2013 via http://www.independent.co.uk/life-style/health-and-families/health-news/drugs-companies-putting-profit-ahead-of-medical-discoveries-warn-scientists-8015784.html.

LePage, M. (2013, August 3). 3D printing has mass appeal. *The Ottawa Citizen*, Business & Technology, pp. D1-D2.

Schlick, T. (2002). 1.1.2 What is Molecular Modeling? In *Molecular modeling and simulation: An interdisciplinary guide*. New York: Springer.

Thompson, M. (2011, November 22). Antibiotic crisis grows while drug companies make lifestyle meds. Retrieved September 8 2013, from http://theconversation.com/antibiotic-crisis-grows-while-drug-companies-make-lifestyle-meds-4373

Appendix I: Figures

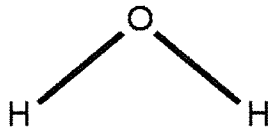

Figure 1: A 2D Chemical Drawing of Water (*Water Chemical Drawing* [Illustration]. (n.d.). Retrieved September 8 2013 from http://faculty.clintoncc.suny.edu/faculty/michael.gregory/files/bio%20101/bio%20101%20lectures/chemistry/chemistr.htm)

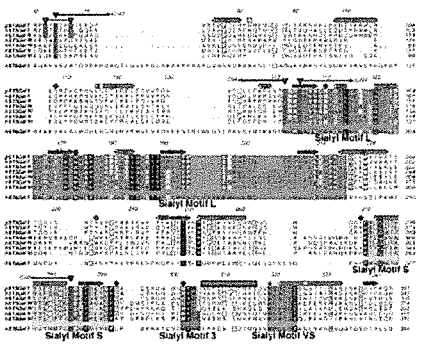

Figure 2: A 2D Sequence Diagram of Human Sialyltransferase (An Enzyme). (Rao, F. V., Rich, J. R., Rakiu0107, B., Buddai, S., Schwartz, M. F., Johnson, K., . . . Strynadka, N. C. J. (2009). Structural insight into mammalian sialyltransferases [PDF]. *Nature Structural and Molecular Biology*, 1-3.)

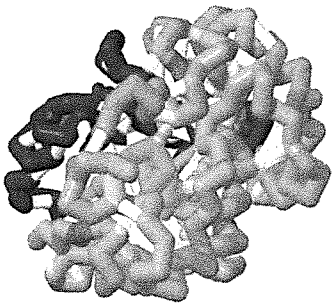

Figure 3: A 3D Digital Model of Human Sialyltransferase (made in Jmol). (Burwash, A. G., Zhang, B., Eid-Ricci, A., & Snoddy, T. (2013). *Human Sialyltransferase (ST3-Gal1)* [Scientific Poster; Microsoft PowerPoint].)

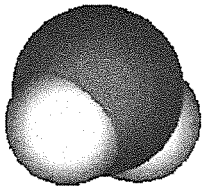

Figure 4: A 3D Digital Model of Water. (*3D Model of Water* [Image]. (n.d.). Retrieved September 8 2013 from http://www.all-water.org/Chemistry.html)

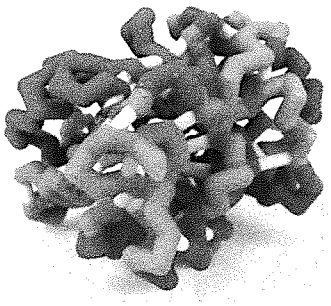

Figure 5: A 3D Printed Model of Human Sialyltransferase. (Burwash, A. G., Zhang, B., Eid-Ricci, A., & Snoddy, T. (2013). *Human Sialyltransferase (ST3-Gal1)* [Scientific Poster; Microsoft PowerPoint].)

Appendix II: Examples of Command Line Programming

This appendix includes some examples of command line programming for 3D models. These commands are some of the multitude of commands that can be used in the modelling program Jmol, a free open source Java based program.

Table 1: Examples of Command Line Programming (*Jmol Quick Reference Sheet* [PDF]. (n.d.). This file is part of the material provided to members of the SMART Team program.)

Desired Action	**Command**
Select everything	Select all
Change the colour of what is selected to green	Colour green
Show the molecule in spacefill view	Spacefill 1.25
Show hydrogen bonds	Hbonds 1.0
Show disulphide bridge	Ssbonds on
Select amino acid 303 in the chain	Select 303

Appendix III: Reflection Journal

American Society of Biochemistry and Molecular Biology Conference April 18-21 2013:

During this conference, after seeing the various presentations that were being made at the conference, I made the decision to move my topic to the world studies area and look specifically at the different types of computerized biological models available and ideas for future models. I did this partly because of the fact that my original topic was extremely specific, and would have therefore been difficult to find relevant research. The second reason why I chose to broaden my topic is due to the fact that the lab portion would not have worked easily in our school's laboratory. While presenting at the undergraduate poster session at the ASBMB conference, I found myself constantly referring to improvements that could be made to the modelling technique, and ways in which current modelling techniques are improvements over previous versions. This included moving from text based studies to 2D models, then 3D models, and hope to eventually move to 3D video models.

Carleton University EE Research Workshop June 17-20 2013:

During this workshop, I found that there are very few journal articles and books published on my topic, since they focus mostly on the program that they are using, rather than the history of the development of these programs. There is, however, plenty of research available on regular websites about the development of the programs and the benefits of each.

After, I searched for information about one of the main programs used for 3D modelling, Jmol. This gave me information about how this particular program has developed, and therefore gave me information about the main parties involved in the evolution of a biological modelling program.

With help from the university librarian, I was able to find books specifically focused on molecular modelling, and they speak of how there is more and more technological modelling coming into play and replacing traditional pen and paper modelling.

I am trying to contact a student at Harvard who has started a database that features different forms of molecular modelling.

I have found a book which has a chapter that goes through the processes that have occurred in molecular modelling. This is exactly what I have been looking for, and it represents a breakthrough for me!

As I keep doing reading on this subject, I keep finding more background information that will be beneficial to me when I continue my work on the SMART team, and will be able to use the information that I gain from my extended essay to increase my knowledge in other areas of my school life.